

Kolkiewicz, Adam; Rice, Gregory; Xie, Yijun

Projection pursuit based tests of normality with functional data. (English) Zbl 1455.62231
J. Stat. Plann. Inference 211, 326-339 (2021).

Summary: Methods for validating the assumption of normality of functional data have been only lightly developed to date, with existing methods based primarily on summarizing the data by their projections into random or principal component subspaces, and applying multivariate normality tests to the vectors of scores defining these projections. While this is effective in some cases, we show with both real and synthetic data examples some pitfalls of this approach, including their sensitivity to the basis used to smooth the raw data. We propose a new normality test for functional data based on a projection pursuit that overcomes some of these challenges. Asymptotic theory is developed for the proposed statistics, and we develop several new computational tools needed to implement the high-dimensional projection pursuit. As a by-product of evaluating the test statistic, our method furnishes a way of decomposing functional data into its approximately Gaussian and non-Gaussian components, which is useful for the purpose of data visualization and subsequent analyses. A simulation study and analysis of three data sets demonstrate the complimentary advantages of the proposed test to those currently available in the literature.

MSC:

[62R10](#) Functional data analysis
[62G10](#) Nonparametric hypothesis testing

Cited in 1 Document

Keywords:

[functional data analysis](#); [projection pursuit](#); [goodness-of-fit](#)

Software:

[LBFGS-B](#); [fda \(R\)](#); [Rainbow](#); [fChange](#)

Full Text: [DOI](#)

References:

- [1] Bali, J. L.; Boente, G.; Tyler, D. E.; Wang, J.-L., Robust functional principal components: A projection-pursuit approach, *Ann. Statist.*, 39, 6, 2852-2882 (2011) · [Zbl 1246.62145](#)
- [2] Baringhaus, L.; Henze, N., Limit distributions for measures of multivariate skewness and kurtosis based on projections, *J. Multivariate Anal.*, 38, 1, 51-69 (1991) · [Zbl 0728.62023](#)
- [3] Benjamini, I.; Lee, S., Conditioned diffusions which are brownian bridges, *J. Theoret. Probab.*, 10, 3, 733-736 (1997) · [Zbl 0884.60077](#)
- [4] Brauchart, J. S.; Dick, J.; Fang, L., Spatial low-discrepancy sequences, spherical cone discrepancy, and applications in financial modeling, *J. Comput. Appl. Math.*, 286, 28-53 (2015) · [Zbl 1317.65005](#)
- [5] Brody, D. C.; Hughston, L. P.; Macrina, A., Information-based asset pricing, *Int. J. Theor. Appl. Finance*, 11, 01, 107-142 (2008) · [Zbl 1152.91487](#)
- [6] Bugni, F. A.; Hall, P.; Horowitz, J. L.; Neumann, G. R., Goodness-of-fit tests for functional data, *Econom. J.*, 12, s1, 1-18 (2009)
- [7] Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C., A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Comput.*, 16, 5, 1190-1208 (1995) · [Zbl 0836.65080](#)
- [8] Cartea, Á.; Jaimungal, S.; Kinzebulatov, D., Algorithmic trading with learning, *Int. J. Theor. Appl. Finance*, 19, 04, Article 1650028 pp. (2016) · [Zbl 1396.91720](#)
- [9] Charniak, E.; Johnson, M., Coarse-to-fine n-best parsing and maxent discriminative reranking, (Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (2005), Association for Computational Linguistics), 173-180
- [10] Clerc, M., *Particle Swarm Optimization*, Vol. 93 (2010), John Wiley & Sons
- [11] Constantinou, P.; Kokoszka, P.; Reimherr, M., Testing separability of space-time functional processes, *Biometrika*, 104, 2, 425-437 (2017) · [Zbl 1506.62541](#)
- [12] Cuesta-Albertos, J.; del Barrio, E.; Fraiman, R.; Matrán, C., The random projection method in goodness of fit for functional

- data, *Comput. Statist. Data Anal.*, 51, 10, 4814-4831 (2007) · [Zbl 1162.62363](#)
- [13] Cuesta-Albertos, J. A.; Fraiman, R.; Ransford, T., Random projections and goodness-of-fit tests in infinite-dimensional spaces, *Bull. Braz. Math. Soc.*, 37, 4, 477-501 (2006) · [Zbl 1112.62037](#)
- [14] Cuevas, A., A partial overview of the theory of statistics with functional data, *J. Statist. Plann. Inference*, 147, 1-23 (2014) · [Zbl 1278.62012](#)
- [15] Cuevas, A.; Febrero, M.; Fraiman, R., An ANOVA test for functional data, *Comput. Statist. Data Anal.*, 47, 111-122 (2004) · [Zbl 1429.62726](#)
- [16] Fletcher, R.; Reeves, C. M., Function minimization by conjugate gradients, *Comput. J.*, 7, 2, 149-154 (1964) · [Zbl 0132.11701](#)
- [17] Friedman, J. H.; Tukey, J. W., A projection pursuit algorithm for exploratory data analysis, *IEEE Trans. Comput.*, 100, 9, 881-890 (1974) · [Zbl 0284.68079](#)
- [18] Georgieva, A.; Jordanov, I., Global optimization based on novel heuristics, low-discrepancy sequences and genetic algorithms, *European J. Oper. Res.*, 196, 2, 413-422 (2009) · [Zbl 1163.90813](#)
- [19] Górecki, T.; Hörmann, S.; Horváth, L.; Kokoszka, P., Testing normality of functional time series, *J. Time Ser. Anal.*, 39, 4, 471-487 (2018) · [Zbl 1416.62489](#)
- [20] Gromenko, O.; Kokoszka, P.; Sojka, J., Evaluation of the cooling trend in the ionosphere using functional regression with incomplete curves, *Ann. Appl. Stat.*, 11, 2, 898-918 (2017)
- [21] Henze, N., Invariant tests for multivariate normality: a critical review, *Statist. Papers*, 43, 4, 467-506 (2002) · [Zbl 1008.62060](#)
- [22] Henze, N.; Wagner, T., A new approach to the BHEP tests for multivariate normality, *J. Multivariate Anal.*, 62, 1, 1-23 (1997) · [Zbl 0874.62043](#)
- [23] Horváth, L.; Kokoszka, P., *Inference for Functional Data with Applications* (2012), Springer · [Zbl 1279.62017](#)
- [24] Jarque, C. M.; Bera, A. K., Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Econom. Lett.*, 6, 3, 255-259 (1980)
- [25] Karlin, S.; Taylor, H. E., *A Second Course in Stochastic Processes* (1981), Academic Press
- [26] Kimura, S.; Matsumura, K., Improvement of the performances of genetic algorithms by using low-discrepancy sequences, *Trans. Soc. Instrum. Control Eng.*, E-6, 1, 16-25 (2007)
- [27] Kowal, D. R.; Matteson, D. S.; Ruppert, D., A Bayesian multivariate functional dynamic linear model, *J. Amer. Statist. Assoc.*, 112, 518, 733-744 (2017)
- [28] Kowal, D. R.; Matteson, D. S.; Ruppert, D., Functional autoregression for sparsely sampled data, *J. Bus. Econom. Statist.*, 37, 1, 97-109 (2019)
- [29] Kruskal, J. B., Linear transformation of multivariate data to reveal clustering, *Multidimens. Scaling*, 1, 101-115 (1972)
- [30] Leobacher, G.; Pillichshammer, F., *Introduction to Quasi-Monte Carlo Integration and Applications* (2014), Springer · [Zbl 1309.65006](#)
- [31] Liang, J.; Li, R.; Fang, H.; Fang, K.-T., Testing multinormality based on low-dimensional projection, *J. Statist. Plann. Inference*, 86, 1, 129-141 (2000) · [Zbl 0961.65005](#)
- [32] Malkovich, J. F.; Afifi, A., On tests for multivariate normality, *J. Amer. Statist. Assoc.*, 68, 341, 176-179 (1973)
- [33] Mardia, K.; Kent, J.; Bibby, J., *Multivariate Analysis* (1979), Academic Press · [Zbl 0432.62029](#)
- [34] Mecklin, C. J.; Mundfrom, D. J., An appraisal and bibliography of tests for multivariate normality, *Internat. Statist. Rev.*, 72, 1, 123-138 (2004) · [Zbl 1211.62095](#)
- [35] Monica, T.; Rajasekhar, A.; Pant, M.; Abraham, A., Enhancing the local exploration capabilities of artificial bee colony using low discrepancy sobol sequence, (*International Conference on Contemporary Computing* (2011), Springer), 158-168
- [36] Niederreiter, H., (*Random Number Generation and Quasi-Monte Carlo Methods. Random Number Generation and Quasi-Monte Carlo Methods, CBMS-NSF Series in Applied Mathematics, vol. 63* (1992), SIAM: SIAM Philadelphia) · [Zbl 0761.65002](#)
- [37] Panaretos, V. M.; Kraus, D.; Maddocks, J. H., Second-order comparison of gaussian random functions and the geometry of DNA minicircles, *J. Amer. Statist. Assoc.*, 105, 670-682 (2010) · [Zbl 1392.62162](#)
- [38] Pant, M.; Thangaraj, R.; Grosan, C.; Abraham, A., Improved particle swarm optimization with low-discrepancy sequences, (*2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)* (2008), IEEE), 3011-3018
- [39] Pedersoli, M.; Vedaldi, A.; Gonzalez, J.; Roca, X., A coarse-to-fine approach for fast deformable object detection, *Pattern Recognit.*, 48, 5, 1844-1853 (2015)
- [40] Ramsay, J. O.; Silverman, B. W., *Applied Functional Data Analysis* (2002), Springer, New York · [Zbl 1011.62002](#)
- [41] Riesz, F.; Sz.-Nagy, B., *Functional Analysis* (1990), Dover
- [42] Shang, H. L.; Hyndman, R. J., *Rainbow: rainbow plots, bagplots and boxplots for functional data* (2016), R package version 3.4
- [43] Sonmez, O.; Aue, A.; Rice, G., *Fchange: Change point analysis in functional data* (2018), R package version 0.2.0
- [44] Szekely, G. J.; Rizzo, M. L., A new test for multivariate normality, *J. Multivariate Anal.*, 93, 1, 58-80 (2005) · [Zbl 1087.62070](#)
- [45] Thadewald, T.; Büning, H., Jarque-bera test and its competitors for testing normality-a power comparison, *J. Appl. Stat.*, 34, 1, 87-105 (2007) · [Zbl 1119.62334](#)
- [46] Yao, F.; Müller, H.-G.; Wang, J.-L., Functional data analysis for sparse longitudinal data, *J. Amer. Statist. Assoc.*, 100,

577-590 (2005) · [Zbl 1117.62451](#)

- [47] Zhu, L.; Fang, K.; Zhang, J., A projection NT-type test for spherical symmetry of a multivariate distribution, *New Trends Probab. Statist.*, 3, 109-122 (1995) · [Zbl 0925.62201](#)
- [48] Zhu, L.-X.; Wong, H. L.; Fang, K.-T., A test for multivariate normality based on sample entropy and projection pursuit, *J. Statist. Plan. Inference*, 45, 3, 373-385 (1995) · [Zbl 0822.62050](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.